

METHODS AND APPARATUS FOR ADAPTIVE PCM LEVEL ESTIMATION AND CONSTELLATION TRAINING

Inventors: Xuming Zhang
Sverrir Olafsson

5

CROSS-REFERENCES TO RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application No.60/118,397, filed February 2, 1999, hereby incorporated by reference.

TECHNICAL FIELD

10 The present invention relates, generally, to modem technology and, more particularly, to adaptive constellation techniques for improving the performance of pulse code modulation (PCM) modems using adaptive constellation design.

BACKGROUND ART AND TECHNICAL PROBLEMS

15 Modem systems operating at about 56 kbps were developed to take advantage of the fact that ISPs are connected to the PSTN through a digital line rather than a twisted pair of copper wires terminating at the central office. That is, referring now to FIG. 1, an ISP 120 generally comprises a server 102 coupled to a digital modem 144 which is connected to PSTN 108 through a digital line 134 and is capable of transmitting data at
20 about 64 kbps. Data is transmitted to central office 110 over a digital line 132, and then over the analog local loop 134 to an analog modem 112 associated with a user system 114 (e.g., a PC or the like).

 PSTN 108 comprises assorted networks and components used to provide, among other things, standard telephone service. PSTN 108 might include foreign exchange
25 services (FX), local exchange carriers (LECs), inter-exchange carriers (IECs), digital loop carriers (DLCs) and the like. As the parameters of PSTN 108 and a line card coder/decoder (codec) provided within central office 110 are dictated by network specifications (e.g., the use of μ -law or A-law encoding), ISP modem 144 is configured to transmit digital data in such a way as to fully exploit its digital connection to the

network. See, e.g., the ITU-T V.90 specification, hereby incorporated by reference.

Digital communication systems may employ a number of initialization, training, and adaptive learning protocols that are designed to equalize the channel distortions, optimize the data transmission speed, reduce transmission errors, and improve the quality of the received signal. For example, the current generation of pulse code modulation (PCM) modems, e.g., modem systems compliant with ITU-T Recommendation V.90, perform an initial training procedure to adaptively adjust the equalizer structure resident at client-side analog modem 112 (APCM).

V.90 modem systems perform an initial two-point training procedure during which one constellation signal point (based on a particular μ -law or A-law level) is transmitted as a sequence having positive and negative signs. The DPCM transmits the two-point training sequence to the APCM, and the APCM analyzes the received signal to determine the channel characteristics and to adjust its equalizers. After performing this two-point training, a digital impairment learning (DIL) procedure is performed.

FIG. 2 illustrates a typical V.90 modem downstream transmission channel 200 over which such a training sequence may be sent. The signal $b(n)$ (202) may represent a sequence of digital symbols, e.g., 8-bit codewords, that are to be transmitted by a DPCM transmitter 204, where " n " represents the time index for the transmitted symbol.

A number of digital impairments 208, such as robbed bit signaling (RBS) and digital pads, may be present within the digital network channel associated with DPCM 204. A digital to analog conversion occurs at a PCM codec 212 to facilitate transmission to the end user over an analog loop as described above. Analog impairments 216, such as nonlinear and linear distortion, may be associated with the analog loop and/or any number of analog processing components. Furthermore, in practical applications, additive noise 218 may be introduced to the analog signal before the analog signal is received by the APCM receiver 220, which produces a series of estimated symbols $b(n)$ (222).

Digital impairments significantly limit the performance of PCM modem receiver 220. Digital impairment includes, for example, digital pads and robbed-bit signaling (RBS). Since reliable operation of a PCM modem is predicated on PCM receiver 220 knowing which levels digital transmitter 204 is sending out, PCM receiver 220 must detect what type of digital impairment has been encountered on a particular telephone line, or estimate all available PCM levels. In typical V.90 systems, RBS-altered symbols are

periodic in nature based on the symbol count; e.g., RBS may occur every six or twelve symbols. Furthermore, the effect of RBS is deterministic but unknown to the APCM, while digital pads cause a constant, deterministic, and level-dependent (nonlinear) effect.

Prior art modem systems may compensate for linear analog impairments, such as amplitude and phase distortions, with well known linear equalization techniques. Such techniques, however, may not adequately compensate for the presence of digital impairments, and may therefore alter the level associated with the predetermined training point. Consequently, the initial training procedure performed by known V.90 modem systems do not provide the most efficient and effective result.

Methods are therefore needed which overcome these and other limitations of the prior art.

SUMMARY OF THE INVENTION

The present invention provides systems for compensating for the presence of digital impairments using adaptive constellation techniques. In accordance with various aspects of the present invention, actual sample magnitudes are estimated from received samples transmitted during a training period using an algorithm based on minimizing the sum of squares of the estimated errors. After which a suitable constellation is derived. In one embodiment, an optimum least-squares algorithm is employed. In an alternate embodiment, a least mean squares (LMS) or normalized LMS algorithm is performed.

BRIEF DESCRIPTION OF THE DRAWING FIGURES

The subject invention will hereinafter be described in conjunction with the appended drawing figures, wherein like numerals denote like elements, and:

FIG. 1 shows a system in which a standard 56K modem may be employed;

FIG. 2 shows a block diagram of a typical V.90 modem downstream transmission channel;

FIG. 3 shows an exemplary ADPCM receiver in accordance with the present invention;

FIG. 4 shows an exemplary FSE+DFE equalizer block-diagram;

FIG. 5 shows an exemplary FSE+DFE/NPD equalizer block diagram;

FIG. 6 shows an alternate block diagram configuration for that shown in FIG. 5.

FIG. 7 shows a flow chart depicting a method in accordance with one aspect of the present invention.

5 DETAILED DESCRIPTION OF PREFERRED EXEMPLARY EMBODIMENTS

The present invention provides methods and apparatus for adaptive constellation training. In general, as will be described in further detail below, sample magnitudes are estimated from received samples transmitted during a training period using an algorithm based on minimizing the mean squares of estimated level errors.

10 As a preliminary matter, the present invention may be described herein in terms of functional block components and various processing steps. It should be appreciated that such functional blocks may be realized by any number of hardware components configured to perform the specified functions. For example, the present invention may employ various integrated circuit components, e.g., memory elements, digital signal processing
15 elements, look-up tables, and the like, which may carry out a variety of functions under the control of one or more microprocessors or other control devices. In addition, those skilled in the art will appreciate that the present invention may be practiced in any number of data communication contexts and that the modem system described herein is merely one exemplary application for the invention. Further, it should be noted that the present
20 invention may employ any number of conventional techniques for data transmission, training, signal processing and conditioning, and the like. Such general techniques that may be known to those skilled in the art are not described in detail herein.

Overview

As discussed above in the connection with **FIG. 2**, analog and digital impairments
25 affect the characteristics of signals transmitted by V.90 modem systems. The affected analog signal is eventually obtained by APCM receiver 220. During the initial training interval, a predetermined training sequence is transmitted by DPCM transmitter 204 and received and processed by APCM receiver 220 such that APCM receiver 220 can adaptively train its equalizers.

FIG. 3 presents a general block diagram depicting a portion of an exemplary APCM receiver 300 that is configured in accordance with the principles of the present invention, and **FIG. 7** is a flow diagram of process 700 that may be performed by the modem system to provide PCM level estimation at APCM receiver 300. Process 700 may begin with a task 702, which causes DPCM 204 to transmit a training sequence of digital symbols.

In a preferred embodiment, the digital impairment learning sequence (DIL) transmitted during training corresponds to each of the 256 PCM levels, wherein the training sequences are sent in blocks of L_i samples (where L_i is the length of the DIL corresponding to PCM level i .) The block length is preferably a multiple of 6). The training sequence for each block is preferably random (i.e., with respect to the signs of the symbols), although other schemes may be used. The sequence can be repeated for all blocks, or each block may use a different sequence of signs. In the case of an LMS scheme (described below), it is advantageous to include a reference pointer periodically in the transmitted samples. For more information regarding an exemplary digital impairment learning sequence, see the V.90 specification.

While being transmitted, digital impairments such as RBS and digital pads may affect the training sequence, and the training sequence is eventually converted to an analog signal for transmission over the local loop, wherein the digital to analog conversion may be performed by, for example, codec 212 shown in **FIG. 2**.

Due to the nature of the transmission channel, APCM 300 receives an analog representation of the training signal (which is a sequence of digital symbols) transmitted by the DPCM. APCM receiver 300 may include or be associated with a linear analog to digital converter (ADC) 301, which is configured to reverse the digital to analog conversion and mapping of codec 212. ADC 301 may be associated with a PCM codec or any suitable operating element of the data communication system. Accordingly, during a task 704, ADC 301 (or any suitable analog-to-digital converter) preferably converts the received analog signal into an equivalent sequence of finite precision (e.g., 16-bit digital symbols) $r(n)$. This received sequence may be considered to be an impaired sequence of digital symbols, i.e., the converted sequence of digital symbols will likely be affected by RBS and/or digital pads. This sequence of digital symbols may then be equalized by a suitable equalizer structure to produce an equalized output $z(n)$.

The equalizer structure is preferably configured to compensate for transmission channel characteristics such as analog impairments. The equalizer structure may be an adaptive equalizer structure having adjustable parameters, e.g, filter taps, that control the operation of the equalizer structure. In the illustrated embodiment, the equalizer structure comprises a feed-forward equalizer 302 and decision feedback equalizer 306. Such pre-cursive and post-cursive equalization techniques are known to those skilled in the art. In general, the equalizer structure may be realized by any number of filter components. For example, equalizer 204 may include a feedforward equalizer (FFE), a decision feedback equalizer (DFE), a noise predictor (NPD), or the like, or any suitable combination of such components. Such components may be fractionally spaced or T-spaced, depending on the specific application. These, and other equalizer and filter arrangements known to those skilled in the art, may be employed in the context of the present invention. A number of such prior art elements are described in detail in ADAPTIVE FILTER THEORY (3rd ed. 1996) by Simon Haykin, the entire contents of which are incorporated herein by reference.

Referring again to **FIG. 3**, a task 706 causes the equalizer structure to suitably equalize the impaired sequence to obtain a corresponding sequence of equalized digital symbols. Following the equalization task, process 700 proceeds to compensate for the presence of digital impairments (e.g., RBS and digital pads) This compensation preferably begins with a task 708, which downsamples the digital symbols produced by FFE 302 using any convenient downsampling.

In the illustrated embodiment, it is preferred that FFE 302 and DFE 306 comprise various filter taps which have been previously trained using a suitable training procedure, for example, a two-point training procedure as set forth in the V.90 specification. For more information regarding such a procedure, See, e.g., U.S. Application Serial No. 08/969,971, filed Sept. 8, 1998. Thus, at this point in process 700, it is assumed that compensation for analog impairments (216 in **FIG. 2**) has been substantially accomplished.

During normal operation, and during the two-point training sequence, slicer 304 generates an output that best approximates the current input symbol. In an ideal optimized system, the input and the output to slicer 304 are substantially close and, consequently, the decision error $e(n)$ would resemble the channel noise. However, due

to the practical nature of the data communication system, some amount of decision error will be present (e.g., error caused by additive noise or imperfections in the equalization process). The decision error is utilized to adaptively adjust FFE 302 and DFE 306 during an initial training sequence in an attempt to optimize the performance of APCM receiver 300. Demapper 308 functions to demap the symbol $d(n)$ to the resultant bit stream $b(n)$.

Returning now to digital impairments, in practical V.90 modem systems, RBS occurring in the digital portion of the transmission channel periodically affects the least significant bit of transmitted symbols. In most cases, RBS forces the least significant bit of the affected symbol to either zero or one. Thus, an RBS-affected codeword may be received as if it were one level less than the transmitted level. Furthermore, RBS is a periodic phenomena in practical systems, i.e., if the symbol at time index $n = n_0$ is affected, then the symbol at time index $n = n_0 + M$ (where M is equal to, for example, 6, 12, 24, etc.) will also be affected. Thus, RBS affects a transmitted sequence of digital symbols in a periodic manner based on a period of N symbols.

In task 710, sample magnitudes are estimated based on the samples which were downsampled in the previous task. In accordance with various aspects of the present invention, sample magnitudes are, in general, estimated using an algorithm based on minimizing the mean square of the estimated errors. The manner in which this task is performed is described in detail below; however, in one embodiment, an optimum least-squares algorithm is employed. In an alternate embodiment, a least mean squares (LMS) or normalized LMS algorithm is employed.

In current V.90 modem systems that operate in conjunction with practical telecommunication networks, RBS is based on a period of six symbols (or an integer multiple of six symbols). RBS is a deterministic (but unknown to the receiver) digital impairment in that once its presence and period are known, the effect it has on any transmitted symbol can be determined. To this end, task 710 preferably determines estimated sample magnitudes for each of the possible RBS symbol positions, or "phases." As a result, a matrix of values is populated during training.

In this regard, it is instructive to review the manner in which the digital impairment learning (DIL) is performed in an exemplary V.90 system. Section 8.4.1 of the V.90 specification sets forth details related to the DIL descriptor as defined in Section 8.3.1. Briefly, the DIL includes N DIL-segments of length L_c , where the number of levels, N , is

between 0 and 255 inclusive, and where $1 \leq c \leq 8$, and $L_c = (H_c + 1) * 6$ symbols. Eight H_c values are used to calculate the length of the DIL-segments containing training symbols. Eight Ucodes, REF_c (or Ref_i , as designated below) define the PCM codeword used as a reference symbol in the various DIL-segments, where the Ucode is the universal code used to describe both u-law and A-law PCM codewords (Table 1/V.90). A sign pattern (SP) and training pattern (TP) is used for the DIL. An SP bit determines the sign of a transmitted symbol (where 0 is negative, and 1 is positive). A TP bit determines whether the reference symbol or a training symbol (designated B_i below) is transmitted (0 for reference, 1 for training symbol). The LSB of each pattern applies to the first symbol of a DIL-segment, where the pattern length is between 1 and 128, inclusive, for both L_{SP} and L_{TP} .

The patterns are restarted at the beginning of each DIL-segment and repeated independently within segments whose lengths exceed L_{SP} or L_{TP} . The sequence is repeated until either the analog modem causes it to be terminated or a timeout occurs. A set of N Ucodes are used for training symbols, where the first Ucode specifies the training symbol for the first DIL-segment, the second Ucode is used for the second DIL-segment, and so on. Additional details related to the DIL sequence may be found in the V.90 specification.

In accordance with the illustrated embodiment, a task 712 builds a constellation based on the estimated samples determined in task 710. The manner in which the system builds the constellation may vary depending upon the particular goals and constraints that apply. For example, the constellation may be configured in accordance with the "minimum distance" between points, signal-to-noise ratio, and or power limit considerations. These and many other criterion for selecting constellations from a particular symbol subset are well known in the art.

Next, in task 714, the constellation built in task 712 is transmitted to the DPCM (204 in FIG. 2). This task may be performed using any convenient method. In the illustrated embodiment, the constellation information is transmitted to the DPCM in accordance with the V.90 specification. The resultant constellation allows the DPCM to transmit data to the APCM such that the digital impairment is partially eliminated or compensated.

Derivation of Optimum Least-Squares Estimation Algorithm

Having thus given an overview of an exemplary system in accordance with the present invention, a detailed derivation and explanation of various methods for determining estimated sample magnitudes (task 710 in FIG. 7) will now be set forth. As a preliminary matter, it will first be assumed that the analog channel can be compensated reasonably well by a well-designed equalizer (such as FSE+DFE/NPD equalizer). The training sequences for PCM level estimation are to be sent on a group of M samples (where M is, for example, a multiple of 6) for each PCM level. The signs of the training sequence for each block is assumed to be random (or pseudo random). This sequence can be repeated for all blocks or each block may use a different binary sequence. As will be shown, the repetitive sequence may have a computational advantage over a totally random sequence.

Begin with a derivation of the PCM level estimation algorithm based on a FSE+DFE structure as illustrated in FIG. 4, which includes FSE 404, DFE 410, and summers 406 and 408. The received signal $r(t)$ 402 is over-sampled and passed through FSE 404. The output of FSE 404 is down-sampled to T -space samples denoted by $\{y(i)\}$, where T is the symbol length.

In such a case, a number of digital impairment learning (DIL) sequences may be appropriate, in an exemplary embodiment, the following DIL sequence is used:

---	---	Ref_i	Ref_i	Ref_i	Ref_i	Ref_i	Ref_i
	$L_i/6$	Ref_i	Ref_i	Ref_i	Ref_i	Ref_i	Ref_i
	---	\vdots					
	$L_i/6$	B_i^l	B_i^l	B_i^l	B_i^l	B_i^l	B_i^l
		B_i^l	B_i^l	B_i^l	B_i^l	B_i^l	B_i^l
		\vdots					
	---	B_i^l	B_i^l	B_i^l	B_i^l	B_i^l	B_i^l

Where L_i is the length of the DIL segment, L_i^r is the length of the reference point "header," B_i^l is the PCM codeword of the l th level on the i th segment, Ref_i is a known reference value. The transmitter works from left to right, and top to bottom, with respect to order of symbols sent in the illustrated matrix. The known reference value might be, for example, zero, or may correspond to the two-point training symbol. In this setup, the Ref_i

points are not necessary for the optimal least-squares method described below, but may be used for other purposes, e.g., to update timing recovery circuitry or equalizer values. It will be appreciated that these reference points (which by definition are known), can be used as an initial guess for LMS or normalized LMS algorithms.

5 For the least-squares solutions, we start to form the necessary vector or matrix when receiving B_i^t . Specifically, let the DFE 410 taps be $\{b_l\}, l = 1, 2, \dots, L$, the estimated symbols be $\{d^{(i)}\}$, and the estimation errors be $\{e^{(i)}\}$. Since a typical RBS frame contains 6 samples, we may decompose the incoming samples into six branches (like a single-input/multiple-output (SIMO) system) and define the following vectors of interest,

$$10 \quad \mathbf{Y}(k) = [y(6k+5), y(6k+4), \dots, y(6k)]^T,$$

$$\mathbf{E}(k) = [e(6k+5), e(6k+4), \dots, e(6k)]^T,$$

$$\mathbf{C} = [c_0, c_1, c_2, c_3, c_4, c_5]^T$$

where c_l is the magnitude of the l -th sample in a RBS frame. Let $s(i) = \text{sign}[d(i)]$ and define a $6 \times (L+6)$ matrix as:

$$15 \quad \mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & b_1 & \dots & b_l \\ 0 & 0 & 0 & 0 & 1 & b_1 & \dots & b_l & 0 \\ 0 & 0 & 0 & 1 & b_1 & b_2 & \dots & 0 & 0 \\ 0 & 0 & 1 & b_1 & b_2 & \dots & 0 & 0 & 0 \\ 0 & 1 & b_1 & b_2 & b_3 & \dots & 0 & 0 & 0 \\ 1 & b_1 & b_2 & b_3 & \dots & 0 & 0 & 0 & 0 \end{bmatrix}$$

and a $(L+6) \times 6$ sparse matrix as

$$S(k) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & s(6k+5) \\ 0 & 0 & 0 & 0 & s(6k+4) & 0 \\ 0 & 0 & 0 & s(6k+3) & 0 & 0 \\ 0 & 0 & s(6k+2) & 0 & 0 & 0 \\ 0 & s(6k+1) & 0 & 0 & 0 & 0 \\ s(6k) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & s(6k-1) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ s(6k-L) & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Note that the matrix B described above should not be confused with the B_i^t levels described above in connection with the DIL sequence. Continuing, the estimation error vector in a training block can be expressed as

$$E(k) = Y(k) - BS(k)C$$

and the least-squares solution of C can be derived by *minimizing* the sum of squares or

$$J(C) = \sum_{k=1}^K \|E(k)\|^2$$

mean squares of the estimated errors, i.e., where $K=(L_i-(L_i^r-L))/6$.

Note that the method of least squares involves the use of time averages and the solution is given by

$$C = R^{-1}p$$

where

$$R = \frac{1}{K} \sum_{k=1}^K S^T(k) B^T B S(k) \quad \text{and} \quad p = \frac{1}{K} \sum_{k=1}^K S^T(k) B^T Y(k)$$

The estimation result of C depends on the number of samples used in the computation. In some cases, an adaptive algorithm might be desired, such as the *recursive least-squares (RLS) algorithm* or its variations (e.g., QR-based algorithms). The RLS algorithm computes the updated estimate of the C vector at time k upon the arrival of new data

vector, given the least-squares estimate of the PCM level vector at $k-1$. Note also that the RLS algorithm can be viewed as a special case of the Kalman filter.

In general, the calculation of \mathbf{C} involves matrix multiplies and accumulates and an inversion of a 6×6 matrix \mathbf{R} . Since the matrix \mathbf{R} is Toplitz, only upper (or lower) triangle
 5 of the matrix needs to be computed and stored. If a common binary training sequence for each PCM level is used for all blocks (a repetitive sequence), the matrices \mathbf{R} and $\mathbf{S}^T(k)\mathbf{B}^T$ need to be calculated only once for all PCM level estimates because it only depends on the DFE taps and signs of the training sequences. Further, since the matrix $\mathbf{S}(k)$ is a sparse matrix (only 1/6 non-zero elements), \mathbf{R} and \mathbf{p} can be easily computed.

10 A further reduction in computation is possible if a pre-classification procedure is performed on the elements in \mathbf{C} before estimating PCM levels. For example, if there is no RBS present in the channel, i.e., $c_0 = c_i, i = 1, \dots, 5$, then

$$c_0 = \frac{\sum_{l=1}^6 \mathbf{p}_l}{\sum_{i,j=1}^6 \mathbf{R}_{i,j}}$$

If one sample in each RBS frame is robbed, then \mathbf{R} is reduced to a 2×2 matrix and \mathbf{p}
 15 becomes 2×1 vector, and so on. Generally speaking, post-classification of \mathbf{C} may result in a slight performance loss because it ignores correlation between the elements in \mathbf{C} when we estimate them.

Finally, a carefully designed, “good,” binary training sequence may result in a matrix of \mathbf{R} close to diagonal and numerical instability of the algorithm might be minimized.

20 The aforementioned estimation algorithm can be easily extended to a FSE+DFE/NPD structure as shown in FIG. 5, which further includes a noise-predictor 502 (NPD) and summer 504. Let the taps of NPD 414 be denoted by $\{q_i\}_{i=1}^N$ and define a new $(L + N - 1) \times 1$ filter coefficient vector as

$$\mathbf{w} = [\mathbf{b}^T, \mathbf{0}^T]^T + [\mathbf{q}^T, \mathbf{0}^T]^T + [0, \mathbf{b}^T \otimes \mathbf{q}^T]^T$$

where \otimes denotes the vector convolution. Then the equalizer in FIG. 5 can be simplified as in FIG. 6 and the proposed algorithm can be adopted. More particularly, a Z^{-1} block 602, an NPD 502, and a W block 604 are configured as shown.

Derivation of LMS Estimation Algorithm

5 Having thus derived the optimum least-squares solution, an estimation algorithm that provides excellent performance while reducing complexity will now be derived. To derive an LMS algorithm, we may either minimize the squared instantaneous estimation error vector $\|E(k)\|^2$ at frame k (vector LMS) or the squared scalar error element $[e(6k+i)]^2$ at time $(6k+i)$ (scalar LMS).

10 An example DIL sequence for the described LMS method is as follows:

---	Ref _i	Ref _i	Ref _i	Ref _i	Ref _i	Ref _i	
	Ref _i	Ref _i	Ref _i	Ref _i	Ref _i	Ref _i	
	B _i [']	Ref _i	Ref _i	Ref _i	Ref _i	Ref _i	---
	Ref _i	B _i [']	Ref _i	Ref _i	Ref _i	Ref _i	
	Ref _i	Ref _i	B _i [']	Ref _i	Ref _i	Ref _i	Initialize
	Ref _i	Ref _i	Ref _i	B _i [']	Ref _i	Ref _i	(c ₀ , . . . c ₅)
L/6	Ref _i	Ref _i	Ref _i	Ref _i	B _i [']	Ref _i	
	Ref _i	Ref _i	Ref _i	Ref _i	Ref _i	B _i [']	---
	Ref _i	Ref _i	Ref _i	Ref _i	Ref _i	Ref _i	
	B _i [']	B _i [']	B _i [']	B _i [']	B _i [']	B _i [']	
	B _i [']	B _i [']	B _i [']	B _i [']	B _i [']	B _i [']	
	⋮						
---	B _i [']	B _i [']	B _i [']	B _i [']	B _i [']	B _i [']	

Given, a DIL as shown above, If the vector LMS approach is used, the level estimate
15 vector is first updated as:

$$C(k+1) = C(k) + \mu E^T(k) S^T(k) W^T$$

where μ is the step size parameter, and $C(k)$ is the vector of the PCM symbol to be estimated at the k th frame. $C(0)$ is preferably set to an initial value. While $C(0)$ may be set to zero, this choice can lead to unsatisfactory convergence times. A better solution,
20 referring now to the previously shown DIL sequence, is to utilize the first portion of the DIL signal to derive reasonable initial values for $(c_0, c_1, \dots c_5)$. Thus, for example, a value

B_i may be placed sparsely in the sequence as shown to estimate each of the six phases (rows 3-8), after which, starting at row 10, the value of k can be reset to zero. The system then proceeds to estimate the levels in accordance with the initial estimate of the C vector, thus reducing convergence time greatly.

5 Returning to the derivation, if the scalar-LMS is used, the level estimate vector can be adapted as:

$$C(6k+i+1)=C(6k+i) + \mu e(6k+i)[S^T(k)W^T]_i$$

where $[B]_i$ denotes the i th column of the matrix B . The instantaneous error can be calculated using newly available level estimates in a round-robin fashion. In one
10 embodiment, the present inventors have found that the normalized-LMS is particularly suitable as it exhibits minimum change in a Euclidean norm sense with respect to the known estimate $C(n)$ at time n . Additionally, the selection of step-size is less sensitive to the analog channel characteristics.

In sum, the scalar version of normalized LMS is given by:

15
$$C(6k+i+1)=C(6k+i) + \mu e(6k+i)[S^T(k)W^T]_i / \|[S^T(k)W^T]\|^2$$

and the vector version is given by

$$C(k+1)=C(k) + \mu E^T(k) S^T(k)W^T / \|S^T(k)W^T\|^2$$

Where the normalized step-size must satisfy $0 \leq \mu < 2$.

In an alternate embodiment, a mapping procedure is defined whereby a signal x
20 is selected every N samples. Then N samples are derived from this signal for transmission based on a particular rule or function. For example, given a signal x , we may select 4 transmit samples as $(x, x, -x, -x)$. If the signals x are selected randomly, the transmitted signal for this mapping procedure will have spectral nulls at DC and $1/2T$. In the case of $1/T=8$ kHz,
25 the nulls would be at 0 and 4 kHz. This greatly simplifies equalization such that feedback is not required.

To identify impairments, the system first transmits an appropriate equalizer training signal using a mapping procedure as described above, based on randomly

selected signals x and $-x$. A FFE is trained based on this sequence. As the training progresses, samples affected by RBS are identified. Typically, samples are affected by RBS in frame sizes of six. The slicing errors in frame positions affected by RBS are then not used for training to avoid tap noise that is otherwise generated because of the uncertainty of the proper slicer level. Once the equalizer is trained to a suitable level, the update is disabled, and a multilevel signal is transmitted. This multi-level signal is modulated using the same mapping procedure as the equalizer training signal. The symbols are selected in such a way that the receiver may estimate which transmit samples are separable in the receiver. Once the receiver is discarded and a new training signal is transmitted -- where feedback is now used -- taking advantage of the knowledge gained about the constellation.

In yet another embodiment, the feedback solution would involve essentially introducing "new" points in a controlled manner. If a signal point Ref_i is known perfectly, and a DFE/NP equalizer has been trained up using that point, a signal point B can be sent once, producing a sequence such as $\text{Ref}_i, \text{Ref}_i, \text{Ref}_i, \text{Ref}_i, B, \text{Ref}_i, \text{Ref}_i, \text{Ref}_i, \text{Ref}_i$. When B is first received, the system simply passes it to the DFE delay line without slicing, and calculates the error for the NP delay line correspondingly. No equalizer update need be performed. The error in subsequent A 's can then be examined to refine the estimate of B based on what is in the DFE and NP delay lines. After several samples have passed, the system can send B again and repeat the process, gradually reducing the interval between B 's. For example, the system could transmit the pattern that follows (split up into RBS frames):

	Ref _i	Ref _i	Ref _i	Ref _i	Ref _i	Ref _i
	B _i ⁺	Ref _i	Ref _i	Ref _i	Ref _i	Ref _i
	Ref _i	B _i ⁺	Ref _i	Ref _i	Ref _i	Ref _i
	Ref _i	Ref _i	B _i ⁺	Ref _i	Ref _i	Ref _i
5	Ref _i	Ref _i	Ref _i	B _i ⁺	Ref _i	Ref _i
	Ref _i	Ref _i	Ref _i	Ref _i	B _i ⁺	Ref _i
	Ref _i	Ref _i	Ref _i	Ref _i	Ref _i	B _i ⁺
	Ref _i	Ref _i	Ref _i	Ref _i	Ref _i	B _i ⁺
	Ref _i	Ref _i	Ref _i	Ref _i	B _i ⁺	Ref _i
10	Ref _i	Ref _i	Ref _i	B _i ⁺	Ref _i	Ref _i
	Ref _i	Ref _i	B _i ⁺	Ref _i	Ref _i	Ref _i
	Ref _i	B _i ⁺	Ref _i	Ref _i	Ref _i	Ref _i
	B _i ⁺	Ref _i	Ref _i	Ref _i	B _i ⁺	Ref _i
	Ref _i	Ref _i	B _i ⁺	Ref _i	Ref _i	Ref _i
15	B _i ⁺	Ref _i	Ref _i	Ref _i	B _i ⁺	Ref _i
	Ref _i	Ref _i	B _i ⁺	Ref _i	Ref _i	Ref _i
	B _i ⁺	Ref _i	Ref _i	Ref _i	Ref _i	B _i ⁺
	Ref _i	Ref _i	Ref _i	B _i ⁺	Ref _i	Ref _i
	Ref _i	B _i ⁺	Ref _i	Ref _i	B _i ⁺	Ref _i
20	Ref _i	Ref _i	Ref _i	B _i ⁺	Ref _i	Ref _i
	Ref _i	B _i ⁺	Ref _i	Ref _i	B _i ⁺	Ref _i
	Ref _i	Ref _i	B _i ⁺	Ref _i	Ref _i	Ref _i
	B _i ⁺	Ref _i	Ref _i	Ref _i	Ref _i	Ref _i

In this case, there are four occurrences of B in each RBS frame, or a total of 24.
 25 Half would actual use +B, the other half -B. If this were done for each point, the
 length of the training signal would be 1.9 seconds. Afterwhich, all signal points could
 be transmitted randomly to refine the estimates.

The present invention has been described above with reference to a preferred
 embodiment. However, those skilled in the art will recognize that changes and
 30 modifications may be made to the preferred embodiment without departing from the scope
 of the present invention. For example, the various adaptive algorithm expressions may

be implemented in alternate ways depending upon the particular application or in consideration of any number of cost functions associated with the operation of the system. In addition, the techniques described herein may be extended or modified for use with other adaptively trainable components in a digital data communication system. These and
5 other changes or modifications are intended to be included within the scope of the present invention, as expressed in the following claims.